桃叶珊瑚属植物的叶绿体基因组结构特征及系统发育分析

李娟，童家赟，范智超，童毅*
（广州中医药大学 中药学院，广州 510006）

摘　要：为确定桃叶珊瑚属（*Aucuba*）植物叶绿体基因组的结构组成及其序列差异，揭示其属下种间亲缘关系，该研究对桃叶珊瑚（*A. chinensis*）、花叶青木（*A. japonica* var. *variegata*）等6种桃叶珊瑚属植物和丝缨花属植物黄杨叶丝缨花（*Garrya buxifolia*）进行二代测序，利用生物信息学软件对其叶绿体基因组序列进行组装和注释，并进行基本特征分析、序列比较以及系统发育分析。结果表明：（1）桃叶珊瑚属植物叶绿体基因组具典型的环状四分体结构，6条序列全长157 891~158 325 bp，均编码114个基因，包括80个蛋白质编码基因、30个tRNA基因和4个rRNA基因；（2）6种植物叶绿体基因组高频密码子数均为29个，偏好以A/U结尾，确定了这6条序列的最优密码子共100个，包含12个共有的最优密码子；（3）6条叶绿体基因组序列共检测到270条散在重复序列，133条串联重复序列以及412个SSR位点；（4）比较基因组学分析结果表明，该属植物叶绿体基因组序列高度保守；（5）从叶绿体基因组中筛选出10个高变片段；（6）系统发育分析结果支持桃叶珊瑚属为一支持率高的单系，其与丝缨花属关系较近。该研究对象的5种桃叶珊瑚属植物以及1种丝缨花属植物的叶绿体基因组均为首次测序组装，揭示了桃叶珊瑚属及其属下种间的系统发育关系，为桃叶珊瑚属植物的分类鉴定和系统发育提供了参考资料。

关键词：桃叶珊瑚属，叶绿体基因组，序列变异，重复序列，密码子偏好性，系统发育分析

中图分类号：Q943　　　　**文献标识码**：A

The sequence characteristics of chloroplast genomes and phylogenic analysis of genus *Aucuba*

LI Juan, TONG Jiayun, FAN Zhichao, TONG Yi *
(*School of Pharmaceutical sciences, Guangzhou University of Chinese Medicine*, Guangzhou 510006, China)

Abstract: To determine the genetic structure and sequence differences of the chloroplast genome of genus *Aucuba* and to reveal their interspecies relationship, the chloroplast genomes of six *Aucuba* species including *A. chinensis* and *A. japonica* var. *variegata* as well as *Garrya buxifolia* were sequenced using high-throughput sequencing technology. The chloroplast genomes were assembled and annotated using Bioinformatics software. The genome characteristics, sequence variation and phylogeny were analyzed based on the chloroplast genomes. The results were as

基金项目：广州中医药大学“青年英才培养工程”项目（QNYC20190105）；国家自然科学青年科学基金项目（31700166）；2019 年医疗服务与保障能力提升补助资金（中医药事业传承与发展部分）“全国中药资源普查项目”（财社〔2019〕39 号）[Supported by the Young Talents Training Project of Guangzhou University of Chinese Medicine (QNYC20190105); National Natural Science Foundation of China (31700166); 2019 Chinese medicine public health service subsidy special “the Fourth Survey on Chinese Materia Medica Resource” (Finance Society (2019) 39)]。

第一作者：李娟（1998-），硕士研究生，主要研究方向为药用植物分类与系统发育，植物条形码，（E-mail）1298194071@qq.com。

*__通信作者__：童毅，博士，副教授，研究方向为药用植物分类与中药资源学，（E-mail）tongyi@gzucm.edu.cn。

follows: (1) Six *Aucuba* plants chloroplast genomes exhibited typical quadripartite circular structure (i.e., LSC, SSC, IRB and IRA) and were between 157 891-158 325 bp in length; (2) A total of 114 genes were annotated (except the duplicate genes on IRs), including 80 protein-coding genes (CDS), 30 tRNA genes and 4 rRNA genes; (3) 270 interspersed repeats, 133 tandem repeats and 412 SSR loci were detected in the six *Aucuba* chloroplast genomes; (4) A total of 29 high-frequency codons prefer ending in A/U. 100 optimal codons were filtered out in total, 12 of which were shared by six *Aucuba* species; (5) Comparative analysis of six *Aucuba* species reveals a highly conserved chloroplast genome sequence; (6) Ten highly divergent regions were identified based on nucleotide diversity analysis; (7) The phylogenetic trees were constructed using maximum likelihood (ML) and Bayesian (BI) showing a similar topology. The results support that genus *Aucuba* was monophyletic, as a sister to genus *Garrya*. We first sequenced the complete chloroplast genomes of these five *Aucuba* species and one *Garrya* species. The results reveal the interspecies relationships and the taxonomic status of genus *Aucuba*, and provide genetic resources for the classification, identification and phylogeny of the genus *Aucuba*.

Key words: *Aucuba*, chloroplast genome, sequence variation, repeated sequences, codon preference, phylogeny

桃叶珊瑚属（*Aucuba* Thunb.），隶属于丝缨花科（Garryaceae）（Angiosperm Phylogeny Group et al., 2016），为常绿小乔木或灌木。该属植物四季常青，冬季红果，为庭园绿化的优良观叶观果树种（Xiang & Boufford, 2005）。本属部分种类民间药用，主要以根、叶和果实入药，具有清热解毒、祛风除湿、活血化瘀等功效（国家中医药管理局中华本草编委会，1997；江纪武，2005；艾铁民，2013；南京中医药大学，2006）。最先从桃叶珊瑚属植物中发现的桃叶珊瑚苷（Aucubin, AU）具有广泛药理活性，包括抗氧化、抗衰老、抗炎、保肝等（Zeng et al., 2020）。

然而，桃叶珊瑚属植物基础的分类修订和物种划分远未完成，这直接影响今后桃叶珊瑚属植物的化学和药理研究，也给药用桃叶珊瑚新资源的开发、推广、生产指导等带来困难。虽然桃叶珊瑚属的花、果实或全株体态使其很容易与其他类群区别，但由于可用作桃叶珊瑚属属下物种鉴定的形态特征复杂多变（如叶形、叶缘、毛被等），种间缺乏有效区分的形态分类性状，且可能存在广泛的杂交和多倍化事件，导致属内物种划分存在很大困难，该属系统而全面的种间系统发育关系亟待建立（Xiang & Boufford, 2005）。

在植物进化过程中，叶绿体基因组的结构和序列相对保守，编码基因的数量、结构、组成和排列顺序基本一致，极少发生重组等变异，被广泛应用于植物的物种鉴定、系统发育和物种起源研究中（刘昶和黄林芳，2020）。目前，在美国国家生物信息中心（NCBI）上共收录了 7 条桃叶珊瑚属植物叶绿体基因组序列，黄媛基于这 7 条叶绿体基因组的 68 个蛋白质编码基因对桃叶珊瑚属进行系统发育分析，结果支持了丝缨花目、丝缨花科和桃叶珊瑚属的单系性（Huang et al., 2022）。

本研究通过浅层基因组测序技术，以已发表的青木（*A. japonica*）的叶绿体基因组（序列号 NC_058874.1）作为参考，组装注释了花叶青木（*A. japonica* var. *variegata*）、峨眉桃叶珊瑚（*A. omeiensis*）、桃叶珊瑚（*A. chinensis*）、密花桃叶珊瑚（*A. confertiflora*）、纤尾桃叶珊瑚（*A. filicauda*）、窄斑叶珊瑚（*A. albopunctifolia* var. *angustula*）以及黄杨叶丝缨花（*Garrya buxifolia*）的叶绿体基因组。通过对桃叶珊瑚属植物的叶绿体基因组进行序列比较及系统发育分析，拟探讨以下问题：（1）6 种桃叶珊瑚属植物叶绿体基因组序列有何结构特征；（2）这些序列之间有何差异；（3）桃叶珊瑚属的系统发育地位如何，6 种桃叶珊瑚

属植物之间有着怎样的亲缘关系。以期为今后深入研究桃叶珊瑚属植物的分类修订、系统发育、生物地理、分子鉴定和物种进化等研究提供叶绿体基因组信息的参考资料。

1 研究材料

桃叶珊瑚属植物的分子材料主要来源于西南和华南分布地野外采集的新鲜叶片，每个个体采集 2~3 片新鲜健康叶片，分别用湿纸巾和吸水纸擦去表面的污渍和水，剪碎后放入茶包袋，用硅胶进行快速干燥。所有凭证标本均保存在广州中医药大学标本室（GUCM），经该文作者分类鉴定，确定这 6 种桃叶珊瑚属植物分别为峨眉桃叶珊瑚、桃叶珊瑚、密花桃叶珊瑚、花叶青木、纤尾桃叶珊瑚和窄斑叶珊瑚（图 1）。黄杨叶丝缨花的干燥叶片取自加州大学伯克利分校植物园（UC Botanical Garden）栽培植物（分子标本采集号 SZ5417）。青木的叶绿体全基因组序列(序列号 NC_058874.1)从 NCBI 数据库(https://www.ncbi.nlm.nih.gov/)下载，用作叶绿体基因组组装及注释的参考，所有新测序的叶绿体基因组序列以及注释信息均已上传至 NCBI 数据库，并获得登录号。采集信息及 GenBank 登录号见表 1。

表 1 桃叶珊瑚属 6 种植物采集信息及 GenBank 序列号

Table 1 Collected information and GenBank accession numbers of six *Aucuba* plants

植物名 Plant name	凭证标本采集 Voucher specimen	采集地点 Collection location	Genbank 序列号 GenBank accession numbers
峨眉桃叶珊瑚 *Aucuba omeiensis*	李娟 LJ20210329026 Li Juan LJ20210329026	四川峨眉山市峨眉山 Emei Mountain, Emei, Sichuan	OQ348516
桃叶珊瑚 *A. chinensis*	李玉玲 LYL264 Li Yuling LYL264	广东惠州市象头山 Xiangtou Mountain, Huizhou, Guangdong	OQ348513
密花桃叶珊瑚 *A. confertiflora*	童毅 TY20081412 Tong Yi TY20081412	广西那坡县老虎跳大峡谷 Laohutiao Valley, Napo, Guangxi	OQ348514
花叶青木 *A. japonica* var. *variegata*	童毅 TY20080506 Tong Yi TY20080506	广西桂林市桂林植物园 Guilin Botanical Garden, Guilin, Guangxi	OQ348512
纤尾桃叶珊瑚 *A. filicauda*	童毅 TY20080903 Tong Yi TY20080903	广西龙胜县红滩瀑布 Hongtan Waterfall, Longsheng, Guangxi	OQ348515
窄斑叶珊瑚 *A. albopunctifolia* var. *angustula*	李娟 LJ20210329002 Li Juan LJ20210329002	四川峨眉山市峨眉山 Emei Mountain, Emei, Sichuan	OQ362998



A. 峨眉桃叶珊瑚；**B**. 桃叶珊瑚；**C**. 密花桃叶珊瑚；**D**. 花叶青木；**E**. 纤尾桃叶珊瑚；**F**. 窄斑叶珊瑚。 **A**. *Aucuba omeiensis*; **B**. *A. chinensis*; **C**. *A. confertiflora*; **D**. *A. japonica* var. *variegata*; **E**. *A. filicauda*; **F**. *A. albopunctifolia* var. *angustula*

图 1 6 种桃叶珊瑚属植物图

Fig.1 Pictures of six *Aucuba* plants

2 方法

2.1 总基因组 DNA 提取与测序

采用改良的 CTAB 方法（Doyle & Doyle, 1987），从硅胶干燥的叶片中提取桃叶珊瑚属植物的总 DNA。提取总 DNA 后，使用 B-500 超微量分光光度计（上海元析仪器有限公司）以及琼脂糖凝胶电泳检测所提 DNA 的质量和浓度，检测合格的 DNA 测序工作委托武汉华大基因科技有限公司使用 DNBSEQ 测序平台进行二代测序。最终每个样品获得 3 GB 的 clean data。

2.2 叶绿体全基因组的组装、注释以及物理图谱的绘制

使用 GetOrganelle-1.7.3.5.（Jin et al., 2020）软件进行叶绿体基因组拼接组装，k-mer 值设置 65、105、127，线程 t 设置 24，其他命令使用默认参数，最终的组装命令为 get_organelle_from_reads.py -1 sample_1.fastq.gz -2 sample_2.fastq.gz -F embplant pt -o output-plastome -R 10 -t 24 -k 65,105,127。然后，以青木的叶绿体全基因组序列作为参考，利用 PGA-master（Qu et al., 2019）软件对组装出的叶绿体基因组进行注释，使用 Geneious R 9.0.2（Basic, 2012）软件对 PGA-master 的注释结果进行手动校正。最后使用在线软件 OGDraw（Lohse et al., 2013）绘制叶绿体基因组物理图谱。

2.3 叶绿体基因组特征分析

使用 Geneious R 9.0.2 软件分别统计 6 种桃叶珊瑚属叶绿体全基因组序列、两个单拷贝区以及一对反向重复区的长度、各区 GC 含量以及基因注释结果等信息。

2.4 密码子偏好性及最优密码子分析

对 6 种桃叶珊瑚植物叶绿体基因组中的 CDS 序列进行筛选，剔除重复基因以及长度小

于 300 bp 的基因，序列中碱基类型仅包含 A、T、C、G，每条序列均含有起始密码子（ATG）和终止密码子（TAG、TGA 和 TAA），序列中间没有终止密码子。最终每个叶绿体基因组序列均获得 52 条符合条件的 CDS，使用 CodonW（Peden，2005）软件计算这些序列的同义密码子相对使用度（RSCU）以及有效密码子数（ENC），所得数据在 Excel 进行整理，并用 TBtools（Chen et al., 2020）软件绘制热图。密码子中 RSCU 值大于 1 的被确定为高频密码子（Wang et al., 2018）。同时满足高频及高表达条件的密码子被认定为最优密码子（Sharp & Li, 1987）。高表达基因具有较低的 ENC 值，低表达基因具有较高的 ENC 值，按照 ENC 值的大小对每个物种的基因进行排序，选取高低两端各 10%的基因分别作为低表达基因组和高表达基因组，计算高表达基因组和低表达基因组每个密码子的差值ΔRSCU，ΔRSCU 大于 0.08 则定为高表达密码子，筛选出高达表密码子中 RSCU 值大于 1 的高频密码子，即得最优密码子（Sharp & Li, 1987）。

2.5 重复序列分析

通过 REPuter（Stefan et al., 2001）在线软件查找叶绿体基因组的重复序列，包括正向重复序列（Forward Repeats）、反向重复序列（Reverse Repeats）、互补重复序列（Complement Repeats）以及回文序列（Palindromic Repeats）4 种重复类型，参数设定为：重复碱基单元 n≥30bp，汉明距离（Hamming distance）为 3。此外，使用 Tandem Repeats Finder（TRF）（Benson, 1999）串联重复序列查找工具的默认参数检测串联重复序列（Tandem Repeats）。使用 MISA（Beier, et al., 2017）软件的 Perl 脚本检测叶绿体基因组的 SSR，从单核苷酸到六核苷酸的重复次数阈值设定依次为 10、5、4、3、3。即 1 个碱基重复≥10 次、2 个碱基重复≥5 次，3 个碱基重复≥4 次，4、5、6 个碱基重复≥3 次。

2.6 IR 边界的收缩和扩张分析

使用在线软件 IRscope（Ali et al., 2018）对 6 条序列进行 IR 边界收缩和扩张的可视化分析，得到叶绿体基因组的简化示意图。

2.7 序列比较分析

以峨眉桃叶珊瑚作为参考序列，利用在线软件 mVISTA（Frazer et al., 2004）对 6 个叶绿体基因组进行序列一致性比较分析，采用 Shuffle-LAGAN 运行模式，生成 6 个叶绿体基因组差异的可视化结果。

2.8 叶绿体基因组序列共线性分析

基因数量及基因顺序的保守性称为共线性，使用 Geneious R 9.0.2 软件里的 Mauve 插件（Darling et al., 2004），对 6 条序列进行共线性分析，用于可视化该属各物种基因排列顺序的一致性和基因重排与倒位现象。

2.9 核苷酸多态性分析

使用 DnaSP v6.0（Rozas et al., 2017）软件对 6 条序列进行滑动窗口分析，检测叶绿体基因组的高变热点区及其大小，变异位点数等信息，并计算核苷酸多样性（nucleotide diversity, P_i）。将比对好的序列导入 DnaSP v6.0，运行参数设定步长（step size）为 200 bp，窗口长度（window length）为 600 bp，筛选出 P_i 值大于 0.01，且片段长度不小于 150 bp 的片段作为桃叶珊瑚属植物叶绿体基因组高变区域，并根据基因注释的结果确定它们在叶绿体基因组上的位置。

2.10 系统发育分析

为了揭示桃叶珊瑚属的系统发育关系，从 NCBI 下载了丝缨花目下的杜仲科杜仲（序列号 NC_037948），加上本次组装所得的黄杨叶丝缨花（序列号 OQ348517），作为外类群，共计 9 条叶绿体基因组序列构建桃叶珊瑚属的系统发育树。使用 Geneious R 9.0.2 里的 MAFFT（Rozewicki et al, 2019）插件进行多序列比较，切掉序列头尾不一样长的碱基后用于系统发育树的构建。通过最大似然法（ML）和贝叶斯推断法（BI）重建桃叶珊瑚属系统发

育关系。使用 IQtree（Minh et al., 2020）构建 ML 树；使用 JModeltest（Posada, 2020）软件进行比对后核酸序列替代模型的选择，以贝叶斯信息准则（Bayesian information criterion, BIC）为标准选出最佳建树模型，使用 MrBayes（Ronquist & Huelsenbeck, 2003）构建 BI 树。

3 结果与分析

3.1 桃叶珊瑚属 6 种植物叶绿体基因组结构和基本特征

6 条桃叶珊瑚属叶绿体全基因组的物理图谱见图 2，6 条序列均为典型的双链环状 DNA，分子序列全长为 157 891~158 325 bp，其总 GC 含量为 37.7%~37.8%。LSC 区的序列长度为 87 210~87 563 bp，GC 含量为 35.8%~35.9%，SSC 区的序列长度为 18 531 ~18 580 bp, GC 含量为 31.5%~31.6%，IR 区的序列长度为 26 067~26 143 bp，GC 含量为 43.0%~43.1%（表 2）。6 条序列均编码 114 个基因（不含 IR 区重复的基因），包括 80 个蛋白编码基因（CDS）、30 个转运 RNA（tRNA）、4 个核糖体 RNA（rRNA）（表 3）。

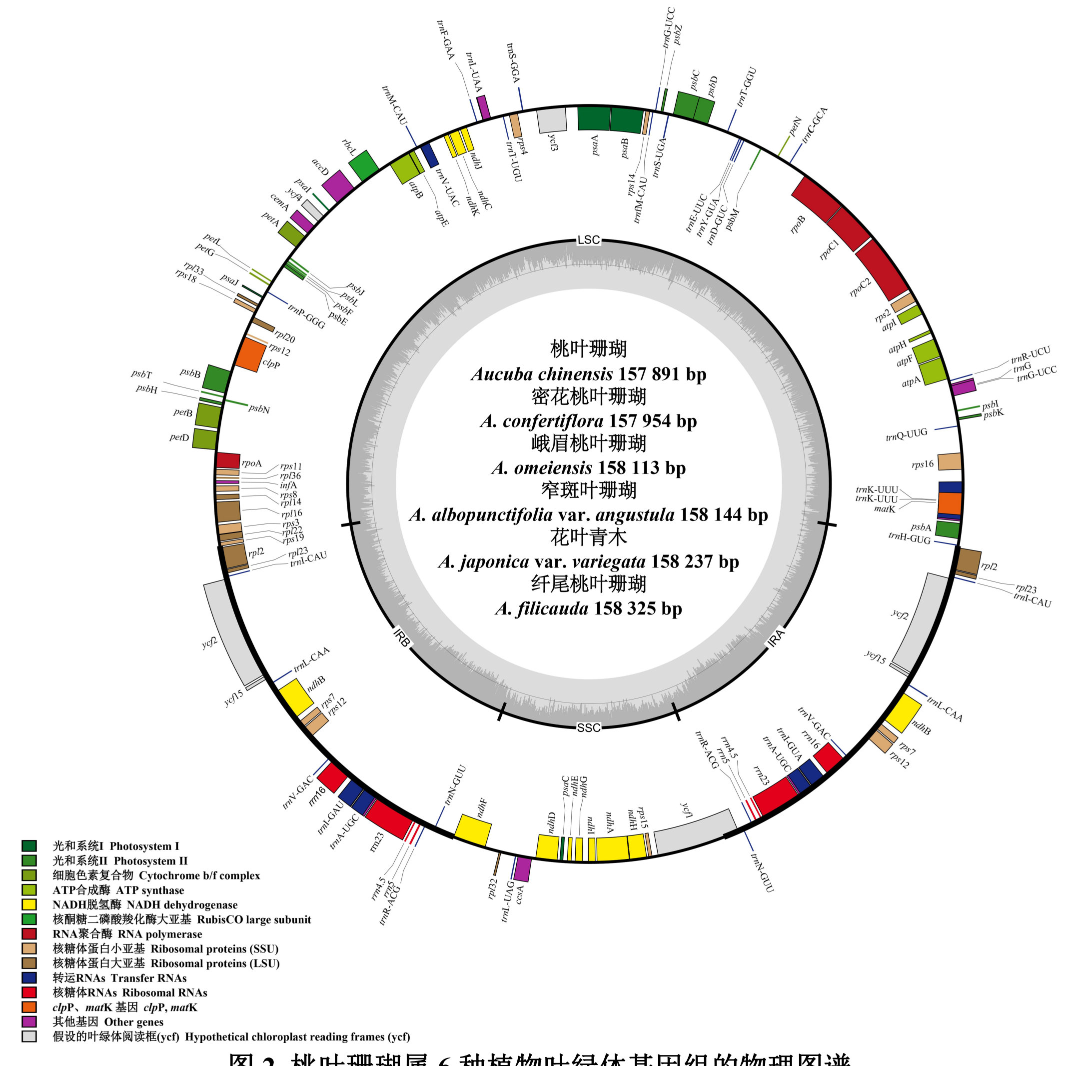


图 2 桃叶珊瑚属 6 种植物叶绿体基因组的物理图谱

Fig.2 Chloroplast genome map of six *Aucuba* plants

表 2 桃叶珊瑚属 6 种植物叶绿体基因组基本特征

Table 2 Chloroplast genomic characteristics of six *Aucuba* plants

物种 Species	峨眉桃叶珊瑚 *Aucuba*	桃叶珊瑚 *A.*	密花桃叶珊瑚 *A.*	花叶青木 *A. japonica* var.	纤尾桃叶珊瑚 *A.*	窄斑叶珊瑚 *A. albopunctifolia*

	omeiensis	*chinensis*	*confertiflora*	*variegata*	*filicauda*	var. *angustula*
基因组大小 Genome size （bp）	158 113	157 891	157 954	158 237	158 325	158 144
大单拷贝区长度 Length of LSC (bp)	87 281	87 210	87 269	87 505	87 563	87 386
大单拷贝区 GC 含量 GC content of LSC (%)	35.9	35.9	35.9	35.9	35.8	35.9
小单拷贝区长度 Length of SSC (bp)	18 546	18 547	18 531	18 544	18 580	18 552
小单拷贝区 GC 含量 GC content of SSC (%)	31.5	31.5	31.5	31.6	31.5	31.5
反向重复区长度 Length of IRs (bp)	26 143	26 067	26 077	26 094	26 091	26 103
反向重复区 GC 含量 GC content of IRs (%)	43.0	43.0	43.1	43.0	43.0	43.0
总 GC 含量 Total GC content (%)	37.7	37.8	37.8	37.7	37.7	37.7
总基因数 Total number of genes	114	114	114	114	114	114
蛋白编码基因数 Number of protein-coding genes	80	80	80	80	80	80
转运 RNA 基因数 Number of tRNA genes	30	30	30	30	30	30
核糖体 RNA 基因数 Number of rRNA genes	4	4	4	4	4	4

表 3 桃叶珊瑚属 6 种植物叶绿体基因组的基因组成

Table 3 Chloroplast genomes composition of six *Aucuba* plants

基因功能 Gene function	基因分类 Gene category	基因名称 Name of gene
自我复制 Self-replication	核糖体蛋白大亚基 Large subunit of ribosome	*rpl*2[a, *]×2、*rpl*14、*rpl*16[*]、*rpl*20、*rpl*22、*rpl*23[a]×2、*rpl*32、*rpl*33、*rpl*36
	核糖体蛋白小亚基 Small subunit of ribosome	*rps*2、*rps*3、*rps*4、*rps*7[a]×2、*rps*8*rps*11、*rps*12[a,b, **]×2、*rps*14、*rps*15、*rps*16[*]、*rps*18、*rps*19[a]
	RNA 聚合酶亚基 DNA-dependent RNA polymerase	*rpo*A、*rpo*B、*rpo*C1[*]、*rpo*C2
	核糖体 RNA Ribosomal RNA	*rrn*4.5[a]×2、*rrn*5[a]×2、*rrn*16[a]×2、*rrn*23[a]×2
	转运 RNA Transfer RNA	*trn*A-UGC[a, *]×2、*trn*C-GCA、*trn*D-GUC、*trn*E-UUC、*trn*F-GAA、*trn*fM-CAU、*trn*G-UCC[*]、*trn*G-UCC、*trn*H-GUG、*trn*I-CAU[a]×2、*trn*I-GAU[a, *]×2、*trn*K-UUU[*]、

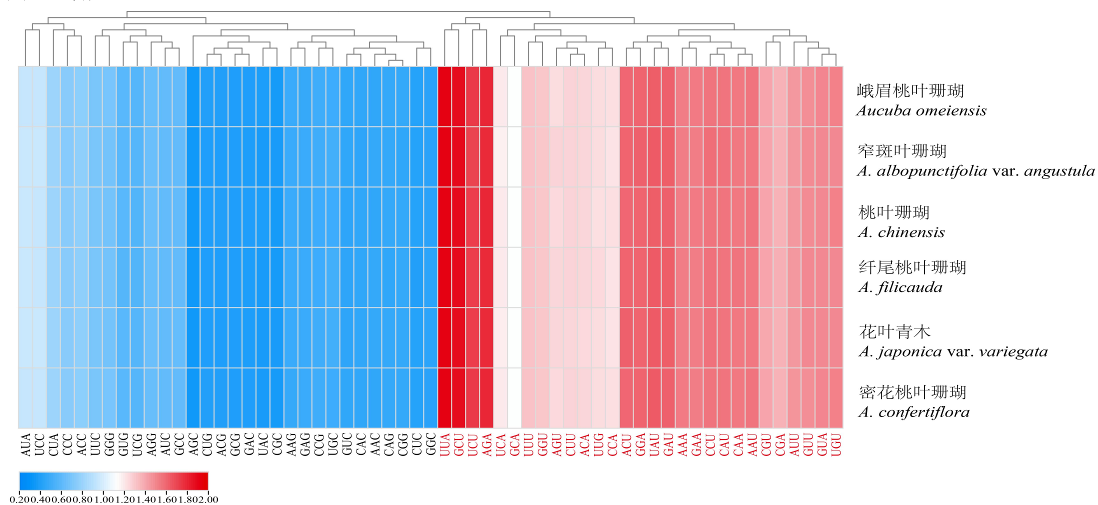
		*trn*L-CAA[a]×2、*trn*L-UAA[*]、*trn*L-UAG、*trn*M-CAU、*trn*N-GUU[a]×2、*trn*P-GGG、*trn*Q-UUG、*trn*R-ACG[a]×2、*trn*R-UCU、*trn*S-GCU、*trn*S-GGA、*trn*S-UGA、*trn*T-GGU、*trn*T-UGU、*trn*V-GAC[a]×2、*trn*V-UAC[*]、*trn*W-CCA、*trn*Y-GUA
光合作用 Photosynthesis	光系统I亚基 Photosystem I	*psa*A、*psa*B、*psa*C、*psa*I、*psa*J
	光系统II亚基 Photosystem II	*psb*A、*psb*B、*psb*C、*psb*D、*psb*E、*psb*F、*psb*H、*psb*I、*psb*J、*psb*K、*psb*L、*psb*M、*psb*N、*psb*T、*psb*Z、*ycf3*[**]
	NADH 脱氢酶亚基 Subunits of NADH-dehydrogenase	*ndh*A[*]、*ndh*B[a, *]×2、*ndh*C、*ndh*D、*ndh*E、*ndh*F[a] *ndh*G、*ndh*H、*ndh*I、*ndh*J、*ndh*K
	细胞色素 b/f 复合物亚基 Subunits of cytochrome b/f complex	*pet*A、*pet*B[*]、*pet*D[*]、*pet*G、*petL*、*petN*
	ATP 合成酶亚基 Subunits of ATP synthase	*atp*A、*atp*B、*atp*E、*atp*F[*]、*atp*H、*atp*I
	核酮糖二磷酸羧化酶大亚基 RubisCO large subunit	*rbcL*
其他基因 Other genes	翻译起始因子 Translation initiation factor	*inf*A
	成熟酶 Maturase	*mat*K
	蛋白酶 Protease	*clpP*[**]
	被膜蛋白 Envelope membrane protein	*cem*A
	乙酰 CoA 羧化酶 Subunits of Acetyl-CoA-carboxylase	*acc*D
	细胞色素 C 合成酶 C-type cytochromesynthesis	*ccs*A
未知功能基因 Unknown function genes	开放阅读框 Open reading frames	*ycf*1、*ycf*2[a]×2、*ycf*4、*ycf*15[a]×2

注：[a]表示位于反向重复区的基因；[b]表示反式剪接基因；[*]表示含有一个内含子的基因；[**]表示含有两个内含子的基因；**×2** 表示双拷贝基因。

Note: [a] indicates genes in inverted repeat regions；[b] indicates trans-splinting gene；[*] indicates genes containing one intron；[**] indicates genes containing two introns；**×2** indicates shows genes with duplicates.

3.2 密码子使用偏好

6 条桃叶珊瑚属叶绿体基因组 CDS 的相对同义密码子使用度（RSCU）见图 3。 RSCU＞1，表明该密码子是偏好密码子，使用较频繁。RSCU＝1 表明该密码子无使用偏好性。RSCU＜1 表明该密码子使用频率较低。在分析时剔除终止密码子 UAA、UGA 和 UAG，以及色氨酸的唯一密码子 UGG 和蛋氨酸的唯一密码子 AUG，因为这些密码子不存在偏好性。结果显示：6 条序列均有 59 个同义密码子，RSCU>1 的高频密码子数均为 29 个，其中，有 28 个以 A 或 U 结尾，有 1 个以 G 结尾，说明桃叶珊瑚属植物叶绿体基因组的密码子偏好以 A 或 U 结尾。



颜色越深表示该密码子的 RSCU 值越高；高频密码子（RSCU>1）已用红色字体标出。

The darker color in the figure indicating the higher RSCU value of the codon; High-frequency codons (RSCU>1) marked in red.

图 3 桃叶珊瑚属 6 种植物叶绿体基因组 RSCU 热图

Fig.3 The RSCU heat map of chloroplast genomes of six *Aucuba* plants

6 条序列共筛选出 100 个最优密码子（表 4），其中，除了窄斑叶珊瑚、花叶青木、纤尾桃叶珊瑚和密花桃叶珊瑚共同拥有的 UUG 以 G 结尾之外，其余最优密码子皆为 A/U 结尾，6 条序列共有的 12 个最优密码子，分别为 AAA、ACU、AGU、CAA、CCU、CGU、GAA、GCU、GGU、GUU、UCU、UGU。其中，峨眉桃叶珊瑚、花叶青木和密花桃叶珊瑚分别特有的最优密码子为 CAU、CUU 和 GAU。

表 4 桃叶珊瑚属 6 种植物叶绿体基因组最优密码子的确定

Table 4 Optimal codons in chloroplast genome of six *Aucuba* species

物种 Species	峨眉桃叶珊瑚 *Aucuba omeiensis*		窄斑叶珊瑚 *A. albopunctifolia* var. *angustula*		桃叶珊瑚 *A. chinensis*		花叶青木 *A. japonica* var. *variegata*		纤尾桃叶珊瑚 *A. filicauda*		密花桃叶珊瑚 *A. confertiflora*	
最优密码子 Optimal codons	ΔRSCU	RSCU	ΔRSCU	RSCU	ΔRSCU	RSCU	ΔRSCU	RSCU	ΔRSCU	RSCU	ΔRSCU	RSCU
AAA	0.42	1.52	0.53	1.52	0.37	1.52	0.44	1.52	0.37	1.52	0.51	1.53
ACU	0.71	1.59	0.63	1.60	0.73	1.59	0.76	1.61	0.89	1.60	0.65	1.59
AGU	0.52	1.23	0.65	1.26	0.52	1.24	0.63	1.25	0.54	1.25	0.79	1.24
AUU	0.14	1.44	0.75	1.44	—	—	0.27	1.44	0.64	1.44	0.72	1.45

CAA	0.27	1.54	0.27	1.54	0.23	1.54	0.31	1.54	0.12	1.55	0.23	1.54
CAU	0.08	1.55	—	—	—	—	—	—	—	—	—	—
CCU	0.93	1.55	0.74	1.57	0.98	1.55	0.95	1.56	0.92	1.56	0.67	1.55
CGA	0.37	1.36	—	—	0.43	1.36	0.51	1.36	—	—	—	—
CGU	0.88	1.40	1.57	1.40	0.88	1.40	0.80	1.39	1.31	1.40	1.57	1.40
CUU	—	—	—	—	—	—	0.13	1.25	—	—	—	—
GAA	0.33	1.53	0.19	1.52	0.30	1.53	0.25	1.52	0.25	1.52	0.18	1.52
GAU	—	—	—	—	—	—	—	—	—	—	0.14	1.61
GCU	0.84	1.87	1.16	1.88	0.80	1.88	1.04	1.87	1.53	1.88	1.29	1.87
GGU	0.81	1.30	1.51	1.30	0.80	1.30	0.73	1.30	1.29	1.30	1.32	1.30
GUA	0.17	1.50	0.51	1.49	0.28	1.49	—	—	0.45	1.49	0.57	1.50
GUU	0.70	1.47	0.81	1.48	0.52	1.48	0.91	1.49	0.47	1.49	0.72	1.47
UAU	0.15	1.62	—	—	0.14	1.62	0.22	1.63	—	—	—	—
UCU	0.78	1.73	0.60	1.70	0.79	1.72	0.73	1.72	0.46	1.72	0.42	1.73
UGU	0.53	1.51	0.91	1.49	0.50	1.51	0.45	1.49	0.89	1.49	0.77	1.51
UUG	—	—	0.56	1.23	—	—	0.12	1.23	0.42	1.24	0.66	1.23
UUA	1.09	1.91	—	—	1.15	1.91	0.77	1.91	—	—	—	—
UUU	0.16	1.31	—	—	0.18	1.31	—	—	—	—	—	—
合计 Total	19		15		17		18		15		16	

注：ΔRSCU=高表达基因 RSCU-低表达基因 RSCU，ΔRSCU≥0.08 即为高表达密码子；RSCU 为桃叶珊瑚属叶绿体基因组 CDS 的相对同义密码子使用度，RSCU＞1 即为高频密码子。

Note: The ΔRSCU is equal to the RSCU of high-expression genes minus that of low-expression genes. If the difference is greater than or equal to 0.08, the codon will be regarded as a highly expressed codon. The RSCU value is the relative synonymous codon usage of *Aucuba* species. If the value of RSCU is greater than 1, the codon will be regarded as a High-frequency codon.

3.3 重复序列和 SSR 分析

利用软件 REPuter 共鉴定出 6 条序列共 270 个散在重复序列，包括 133 个正向重复序列、8 个反向重复序列、2 个互补重复序列，127 个回文重复序列（见表 5）。6 条序列中，正向重复序列最多，回文重复序列次之，反向重复序列和互补重复序列较少存在，仅在桃叶珊瑚和窄斑叶珊瑚中分别检测到了一条互补重复序列，其他 4 种均未检测到。

表 5 桃叶珊瑚属 6 种植物叶绿体基因组中散在重复序列的类型和数量

Table 5 Types and number of interspersed repeats in the chloroplast genomes of six *Aucuba* plants

物种 Species	散在重复序列的数量 Number of interspersed repeats				
	正向重复序列 Forward repeat	反向重复序列 Reverse repeat	互补重复序列 Complement repeat	回文重复序列 Palindromic repeat	合计 Total
密花桃叶珊瑚 *Aucuba confertiflora*	21	1	0	20	42
花叶青木 *A. japonica* var. *variegata*	18	1	0	18	37
纤尾桃叶珊瑚	24	2	0	22	48

A. filicauda					
桃叶珊瑚 *A. chinensis*	21	2	1	20	44
窄斑叶珊瑚 *A. albopunctifolia* var. *angustula*	20	1	1	20	42
峨眉桃叶珊瑚 *A. omeiensis*	29	1	0	27	57

通过 TRF 对串联重复序列进行检测，统计情况见表 6。6 条序列中串联重复序列共 133 个；其中检测到串联重复序列最少的是密花桃叶珊瑚，为 20 个；而花叶青木最多，达 25 个。桃叶珊瑚属植物叶绿体基因组的串联重复序列在 LSC 区和 IR 区的分布明显多于 SSC 区。

表 6 桃叶珊瑚属 6 种植物叶绿体基因组中串联重复序列统计

Table 6 Total counts of tandem repeats in the chloroplast genomes of six *Aucuba* species

物种 Species	不同区域串联重复序列的数量 Number of tandem repeats in different regions				
	大单拷贝区 LSC	反向重复区 B IRB	小单拷贝区 SSC	反向重复区 A IRA	合计 Total
密花桃叶珊瑚 *Aucuba confertiflora*	4	5	5	6	20
花叶青木 *A. japonica* var. *variegata*	11	5	3	6	25
纤尾桃叶珊瑚 *A. filicauda*	6	7	4	7	24
桃叶珊瑚 *A. chinensis*	5	5	5	6	21
窄斑叶珊瑚 *A. albopunctifolia* var. *angustula*	7	6	3	6	22
峨眉桃叶珊瑚 *A. omeiensis*	4	6	5	6	21

MISA 软件检测结果显示，6 条序列共检测到 412 个 SSR 位点，共识别出 4 种 SSR，包括单核苷酸 SSR（367 个）、二核苷酸 SSR（26 个）、三核苷酸 SSR（9 个）、四核苷酸 SSR（10 个）；SSR 的类型以单核苷酸 A/T 重复为主，占比所有 SSR 位点数的 86.88%~99.41%，其中，纤尾桃叶珊瑚和花叶青木被识别出最多的 SSR 个数，均为 73 个；桃叶珊瑚被识别出最少的 SSR 个数，为 61 个（表 7）。

表 7 桃叶珊瑚属 6 种植物叶绿体基因组 SSR 在不同区域数量分布

Table 7 Number and distribution of SSR in the chloroplast genomes of six *Aucuba* species

物种 Species	分布区域 Distribution of SSR	SSR 类型及数量 Type and number of SSR											
		A	T	C	G	AT	AT TT	T A	TA T	TT C	A C A A	T T A T	合计 Total
密花桃叶珊瑚	大单拷贝区	24	22	0	0	2	0	2	1	1	1	0	53

Aucuba confertiflora	LSC												
	反向重复区 B IRB	0	2	0	1	0	0	0	0	0	0	0	3
	小单拷贝区 SSC	3	3	0	0	1	0	0	0	0	0	0	7
	反向重复区 A IRA	2	0	1	0	0	0	0	0	0	0	0	3
花叶青木 *A. japonica* var. *variegata*	大单拷贝区 LSC	25	23	1	0	2	0	1	0	1	1	1	55
	反向重复区 B IRB	0	3	0	1	0	0	0	0	0	0	0	4
	小单拷贝区 SSC	4	5	0	0	1	0	0	0	0	0	0	10
	反向重复区 A IRA	3	0	1	0	0	0	0	0	0	0	0	4
纤尾桃叶珊瑚 *A. filicauda*	大单拷贝区 LSC	26	25	0	0	2	0	1	0	1	1	1	57
	反向重复区 B IRB	0	3	0	0	0	0	0	0	0	0	0	3
	小单拷贝区 SSC	4	5	0	0	1	0	0	0	0	0	0	10
	反向重复区 A IRA	3	0	0	0	0	0	0	0	0	0	0	3
桃叶珊瑚 *A. chinensis*	大单拷贝区 LSC	22	22	0	0	2	0	1	1	1	1	0	50
	反向重复区 B IRB	0	2	0	0	0	0	0	0	0	0	0	2
	小单拷贝区 SSC	1	4	0	0	1	1	0	0	0	0	0	7
	反向重复区 A IRA	2	0	0	0	0	0	0	0	0	0	0	2
窄斑叶珊瑚 *A. albopunctifolia* var. *angustula*	大单拷贝区 LSC	24	26	0	0	2	0	1	0	1	1	1	56
	反向重复区 B IRB	0	3	0	0	0	0	0	0	0	0	0	3
	小单拷贝区 SSC	3	5	0	0	1	0	0	0	0	0	0	9
	反向重复区 A IRA	3	0	0	0	0	0	0	0	0	0	0	3
峨眉桃叶珊瑚 *A. omeiensis*	大单拷贝区 LSC	24	23	1	0	2	0	2	1	1	1	0	55
	反向重复区 B IRB	0	2	0	1	0	0	0	0	0	0	0	3
	小单拷贝区	3	3	0	0	1	0	0	0	0	0	0	7

SSC												
反向重复区 A IRA	2	0	1	0	0	0	0	0	0	0	0	3
合计 Total	178	181	5	3	18	1	8	3	6	6	3	412

3.4 IR/SC 边界的扩张与收缩分析

桃叶珊瑚属植物叶绿体全基因组的 2 个 IR 区在 LSC 及 SSC 之间存在 LSC/IRB、IRB/SSC、SSC/IRA 以及 IRA/LSC4 这 4 个边界（图 4）。4 个边界相对保守，但各种之间仍具较小差异，LSC/IRB 的边界均位于 *rps*19 基因内部，6 条序列的 *rps*19 基因均向 IRB 区扩张了 33 bp；SSC/IRB 边界均位于 *ndh*F 基因内部，6 条序列的 *ndh*F 基因均向 IRB 区扩张了 42 bp；SSC/IRA 边界均位于 *ycf*1 基因内部，密花桃叶珊瑚和花叶青木的 *ycf*1 基因向 IRA 区扩张了 1 082 bp，纤尾桃叶珊瑚、窄斑叶珊瑚和峨眉桃叶珊瑚的 *ycf*1 基因向 IRA 区扩张了 1 079 bp，桃叶珊瑚的 *ycf*1 基因向 IRA 区扩张了 1 061 bp；LSC/IRA 边界位于 *trn*H 附近，峨眉桃叶珊瑚与密花桃叶珊瑚的 LSC/IRA 边界与 *trn*H 的距离为 14 bp，其余 4 种的 LSC/IRA 边界与 *trn*H 的距离均为 7 bp。

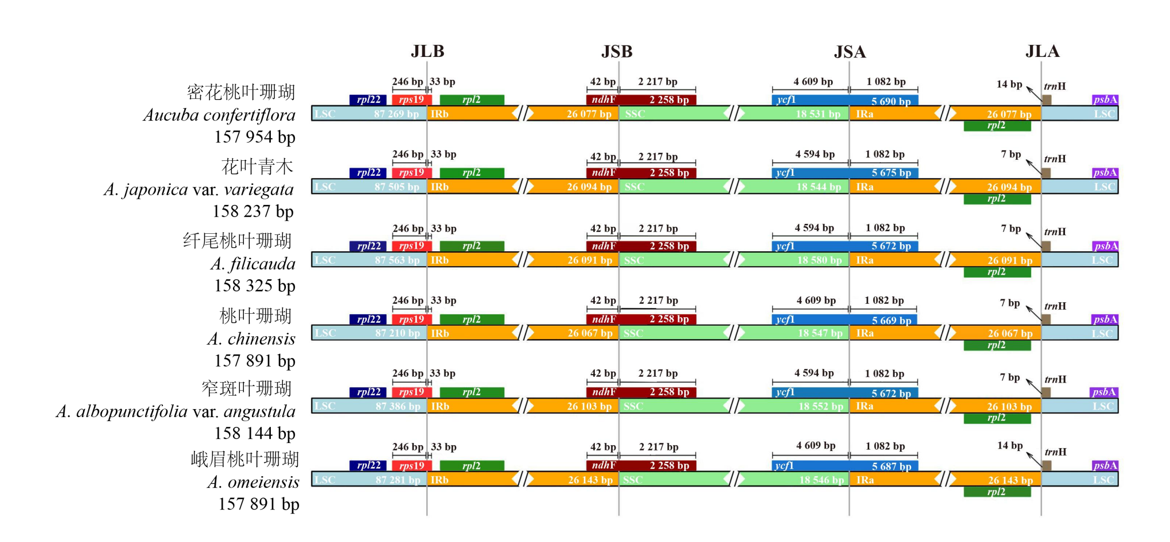
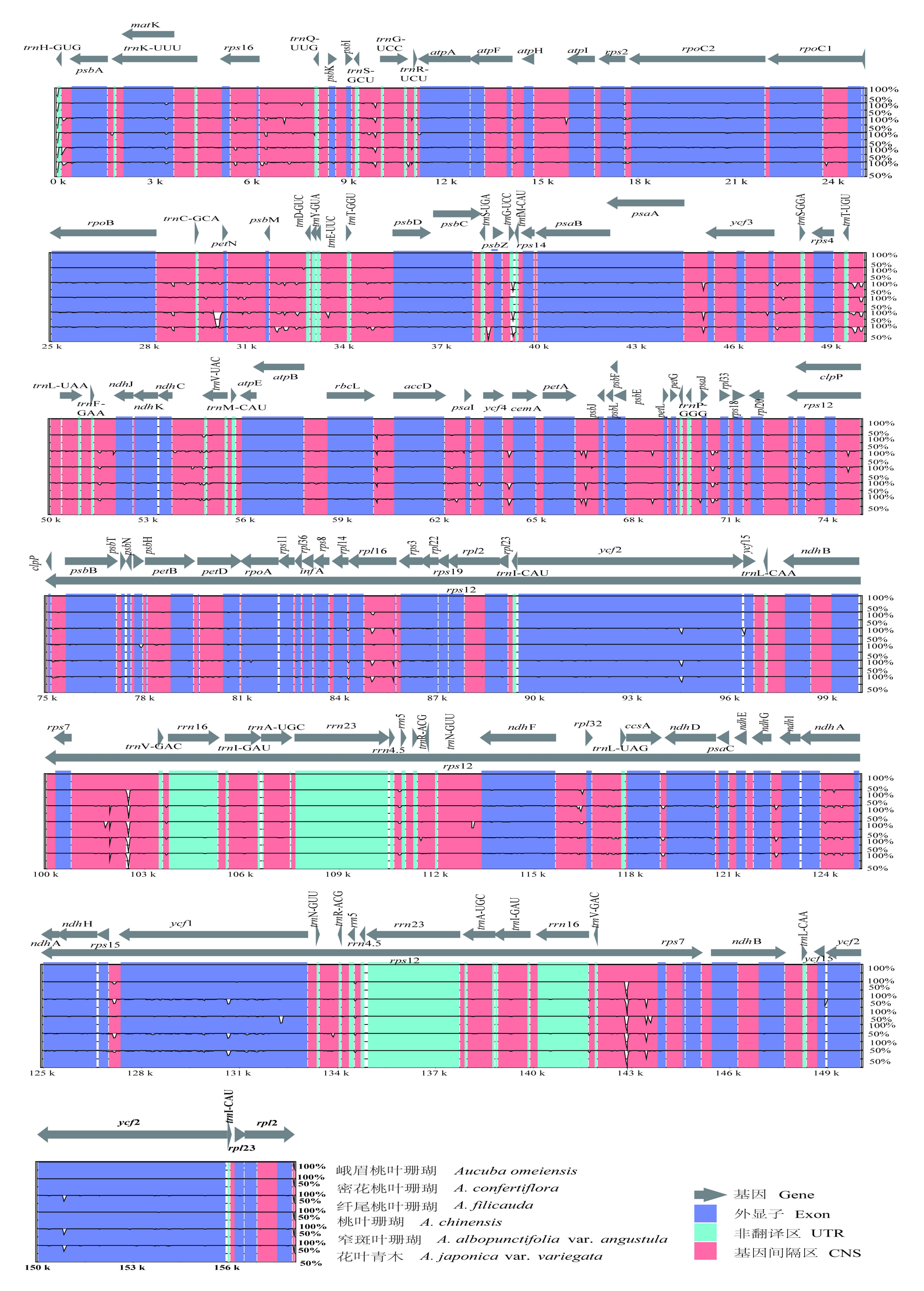


图 4 桃叶珊瑚属 6 种植物叶绿体基因组 LSC、IR 和 SSC 边界比较示意图

Fig.4 Comparison of the LSC, IR and SSC boundaries in the chloroplast genomes of six *Aucuba* plants

3.5 序列变异分析

为了比较桃叶珊瑚属内物种间叶绿体基因组序列的差异，以峨眉桃叶珊瑚的叶绿体基因组序列为参考，将该 6 条序列进行全局对比，mVISTA 图的结果显示了桃叶珊瑚属 6 条序列的高度相似性，单拷贝区比反向重复区更保守，编码区比非编码区（non-coding sequences, NCS）相对更保守（图 5）。位于 LSC 区的 *trn*C-GCA- *pet*N 以及位于 IR 区的 *rps*7- *trn*V-GAC 均为基因间隔区，在这些区域相似度较低，存在不同程度的变异。此外，rRNA 基因和 tRNA 基因区域的一致度最高，序列最为保守。LSC 和 SSC 区序列相比于 IR 区有更多的变异，表明 IR 区在进化过程中更为保守。



灰色箭头表示基因的方向；深蓝色区域表示编码区；红色区域表示非编码序列；浅蓝色区域表示 tRNA 或 rRNA；*Y* 轴表示从 50%到 100%序列之间的相似性。

Gray arrows and thick black lines indicates gene orientation; Dark blue indicates exons; Red regions indicates non-coding Sequences; Light-blue indicates tRNA or rRNA; *Y* axis indicates the similarity among sequences, ranging from 50% to 100%.

图 5 桃叶珊瑚属 6 种植物叶绿体基因组序列可视化比对分析

Fig.5 Plastomes comparison of six *Aucuba* species

3.6 叶绿体基因组共线性分析

叶绿体基因组的多重基因组比对法检测出 6 条序列之间只有一个局部共线区（locally collinear blocks, LCBs）（图 6），所有基因的种类、数量和排列顺序在属内均高度一致，叶绿体基因组完全共线，没有重排重组现象发生，进一步说明该属叶绿体基因组具有较高的保守性。

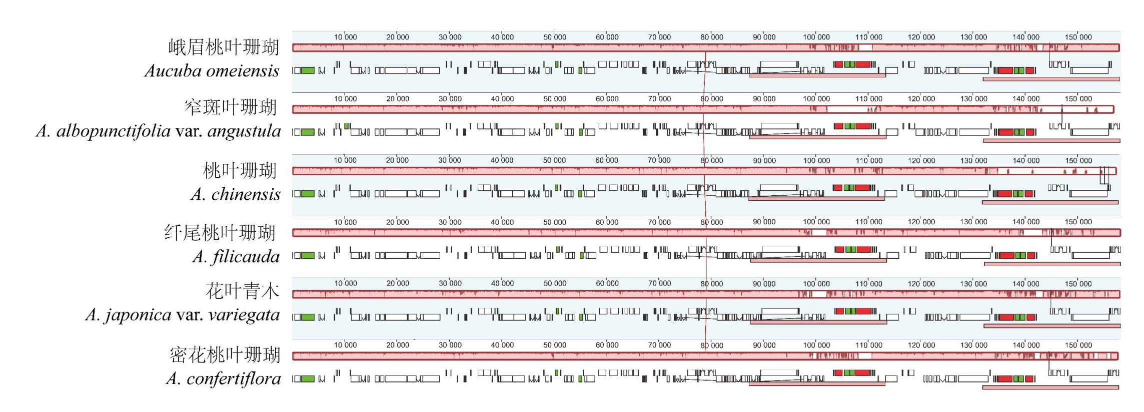


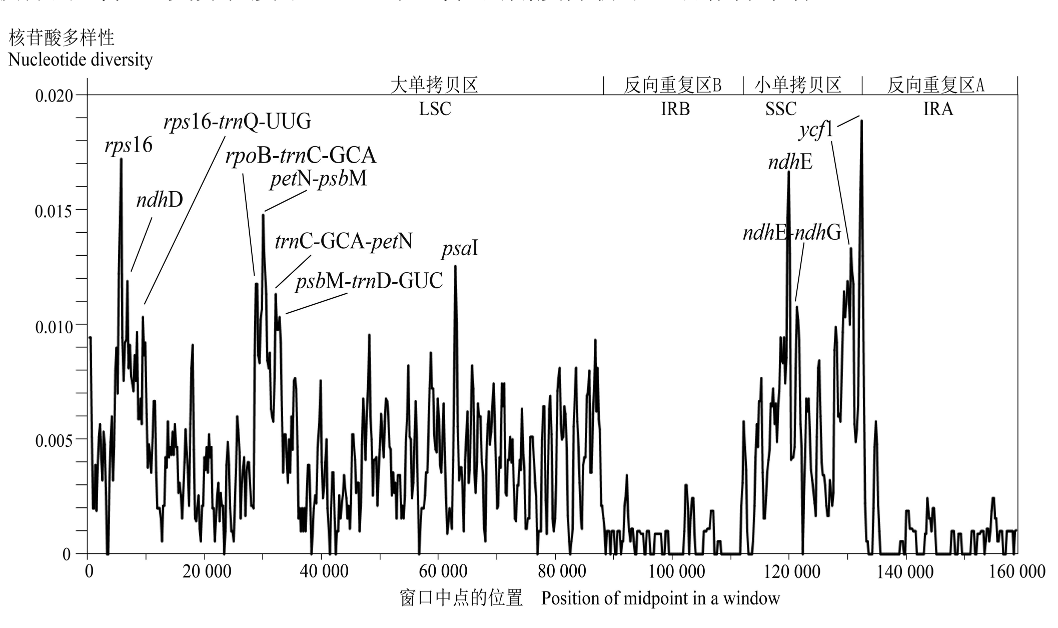
图 6 桃叶珊瑚属 6 种植物叶绿体全基因组 Mauve 共线性分析

Fig.6 Mauve genome alignments of the whole chloroplast genomes of the six *Aucuba* plants

3.7 序列变异热点分析

滑窗分析结果显示，6 条序列的核苷酸多样性（P_i）在 0~0.018 89 之间，平均 P_i 值为 0.003 51（图 7），表明 6 种植物的叶绿体基因组高度相似，序列非常保守；其 IR 区的核苷酸多样性（P_i=0.000 85）明显低于 LSC 区（P_i=0.004 47）和 SSC（P_i=0.006 14）区，SSC 显示出最高的核苷酸多样性，而 IR 区在整个叶绿体基因组中更为保守。共筛选出 10 个高变片段（表 8），包括位于 LSC 区的 *rps*16、*rps*16-*trn*Q-UUG、*rpo*B-*trn*C-GCA、*pet*N-*psb*M、*trn*C-GCA-*pet*N、*psb*M-*trn*D-GUC 和 *acc*D-*psa*I 七个片段，以及位于 SSC 区的 *ndh*E、*ndh*E-*ndh*G 和 *ycf*1 三个片段，其中，*ycf*1 的分化程度最高。在这些高变区域中，仅有 *rps*16、*ndh*E 和 *ycf*1 三个片段是编码序列，其他高变片段均位于基因间隔区（intergenic spacer，IGS），这些高变片段可作为桃叶珊瑚属植物物种鉴定的条形码候选片段。

目前，叶绿体基因组中被广泛应用于植物鉴定的 3 条通用的条形码为 *mat*K、*rbc*L 和 *trn*H-GUG-*psb*A，但是在本研究中，这 3 个片段在桃叶珊瑚属叶绿体基因组序列中均显示了较低的 P_i 值，变异程度不理想，从 P_i 值的角度分析不适合做分子标记。



窗口长度 600 bp，步长 200 bp。10 个高变区(高 P_i 值)均被标记。

Window lengths is 600 bp; step size is 200 bp. Ten regions with the highest P_i values are marked out.

图 7 基于滑窗分析检测桃叶珊瑚属植物叶绿体全基因组的高变区

Fig.7 Hypervariable regions within the chloroplast genomes of six *Aucuba* species using a sliding window analysis

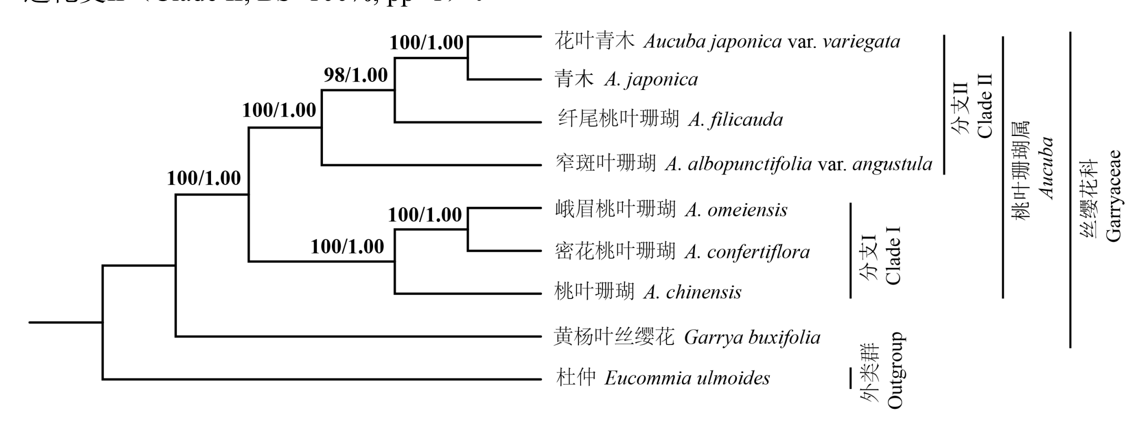
表 8 桃叶珊瑚属 6 种植物叶绿体基因组 10 个高变区信息

Table 8 Ten hypervariable regions of chloroplast genomes of six *Aucuba* plants

序号 Number	片段 Fragment	片段长度 Fragment size (bp)	核苷酸多态性 P_i	分布区域 Distribution region
1	*rps*16	1 132~1 139	0.017 22	LSC
2	*rps*16-*trn*Q-UUG	1 739~1 794	0.012 22	LSC
3	*rpo*B-*trn*C-GCA	1 251~1 267	0.011 78	LSC
4	*pet*N-*psb*M	1 152~1 172	0.014 78	LSC
5	*trn*C-GCA-*pet*N	700~830	0.011 33	SSC
6	*psb*M-*trn*D-GUC	1 172~1 192	0.010 33	LSC
7	*acc*D-*psa*I	685~704	0.012 56	LSC
8	*ndh*E	306	0.016 67	SSC
9	*ndh*E-*ndh*G	227	0.010 78	SSC
10	*ycf*1	5 670~5 691	0.018 89	SSC

3.8 系统发育分析

根据 BIC 方法用 JModeltest 筛选出核苷酸的最佳替代模型均为 GTR+G。分别使用软件 MrBayes 和 IQtree 构建 BI 树和 ML 树，结果见图 8。系统发育分析结果表明，基于叶绿体全基因组构建的 BI 树和 ML 树展现出了完全一致的拓扑结构，桃叶珊瑚属所有物种形成一个支持率高的单系（BS=100%, pp=1），与丝缨花属构成姊妹类群，共同组成丝缨花科。桃叶珊瑚属属内分化为两个进化支，峨眉桃叶珊瑚、密花桃叶珊瑚以及桃叶珊瑚共同组成进化支I（Clade I, BS=100%, pp=1）；花叶青木、青木、纤尾桃叶珊瑚以及窄斑叶珊瑚共同组成进化支II（Clade II, BS=100%, pp=1）。



分支上的数字分别为自展支持率和贝叶斯后验概率。

Numbers above the branches are the bootstrap values and Bayesian posterior probabilities.

图 8 基于丝缨花目 9 条叶绿体全基因组构建的系统发育树（ML 和 BI）

Fig.8 Phylogenetic relationships inferred from maximum likelihood and Bayesian inference based on nine Garryales chloroplast genome datasets (ML and BI)

4 讨论与结论

4.1 叶绿体基因组的结构特征

该研究组装注释了桃叶珊瑚属 6 种植物的叶绿体基因组，并进行了序列比较分析，结果显示 6 条序列是高度相似的，无论是基因组结构及大小，还是基因的含量及组成，皆是相对保守的。桃叶珊瑚属叶绿体基因组拥有典型的环状四分体结构，6 条序列全长 157 891~158 325 bp，6 条序列均编码 114 个基因，包括 80 个蛋白质编码基因、30 个 tRNA 基因以及 4 个 rRNA 基因，总 GC 量为 37.7%~37.8%。LSC、SSC、IR 区的 GC 含量分别为 35.8%~35.9%、31.5%~31.6%、43.0%~43.1%，与大多数被子植物一样，桃叶珊瑚属叶绿体基因组的 IR 区相较于 LSC 区和 SSC 区都更为稳定，有着最高的 GC 含量。

密码子对于遗传信息的正确表达至关重要，密码子的使用偏好在不同物种甚至某一物种内的不同基因都具有特异性，这是由于基因与物种长期进化过程中选择、突变以及漂变等因素的综合作用，亲缘关系越近或者生存环境越相似的物种越有可能采用相似的密码子选择策略（Romero, et al., 2000; Xu, et al., 2011），通过比较密码子偏好性差异可以推断基因是否受到不同程度的翻译选择，对于探讨桃叶珊瑚属物种的进化模式有着重要的作用。6 条序列均有 59 个同义密码子，RSCU>1 的高频密码子数均为 29 个，其中，除了 UUG 以 G 结尾之外，其余 28 个高频密码子均以 A/U 结尾，这表明桃叶珊瑚属叶绿体基因组的密码子偏好以 A/U 结尾；在 6 条序列中共筛选出 100 个最优密码子，包含 12 个共有最优密码子，分别为 AAA、ACU、AGU、CAA、CCU、CGU、GAA、GCU、GGU、GUU、UCU、UGU，均是以 A/U 结尾，尤其是以 U 结尾，这与双子叶植物偏好使用 A/U 结尾密码子的研究结果一致（Kawabe & Miyashita, 2003）。

叶绿体重复序列的存在和丰富程度可增加生物种群遗传多样性，且与多种系统发育信号有关（Adeyemo et al., 2021），SSR 和重复序列在叶绿体基因组中广泛存在，具有丰富的多态性位点，易复制、遗传信息含量高等优点，在研究物种遗传多样性、亲缘关系、品种鉴定和标记辅助育种中具有较高的可靠性（Adeyemo et al., 2021；贾守宁等，2022）。6 条叶绿体基因组序列共检测到 412 个 SSR 位点，SSR 的类型以单核苷酸 A/T 重复为主，且主要位于 LSC 区； 6 条序列共检测到 270 条散在重复序列以及 133 条串联重复序列。该属叶绿体基因组上的重复序列的功能还需进一步的深入研究。

4.2 叶绿体基因组的结构比较分析

比较基因组学分析结果表明，桃叶珊瑚属叶绿体基因组序列具有高度的相似性，单拷贝区比反向重复区更为保守，编码区比非编码区更为保守；叶绿体基因组 IR/SC 边界分布的基因类型完全相同；基因组基因结构和排列顺序整体上相似，所有基因的种类、数量和排列顺序均高度一致，6 条序列完全共线，没有重排重组现象发生；根据课题组前期研究（未发表）发现，目前通用的 DNA 条形码片段（*psb*A-*trn*H，*rbc*L，*mat*K，ITS，ITS2）及其片段组合的变异程度不理想，对桃叶珊瑚属植物的物种分辨率不高，不适合作为该属植物的通用条形码，该文从桃叶珊瑚属叶绿体基因组 LSC 区和 SSC 区共筛选出 10 个变异率较高，长度适中的高变片段，这些高变区可作为桃叶珊瑚属物种鉴定的潜在条形码，可以进一步结合双亲遗传的核基因片段联合分析，为桃叶珊瑚属植物的物种鉴定，杂交起源，多倍体物种的形成和系统进化分析提供可靠的分子标记。

4.3 系统发育分析

系统发育分析结果表明，桃叶珊瑚属为一个支持率高的单系（BS=100%, pp=1），与丝缨花属亲缘关系较近，共同构成丝缨花科；丝缨花科与杜仲科关系较近，这与 APG IV和 Huang 等（2022）的研究结果一致。

桃叶珊瑚属属内分化为两个支持率高的进化支，峨眉桃叶珊瑚、密花桃叶珊瑚以及桃叶

珊瑚共同组成进化支I，共同的特征为乔木、花绿色；花叶青木、青木、纤尾桃叶珊瑚以及窄斑叶珊瑚共同组成进化支II，共同的特征为灌木、花红色。这两组特征是否分别稳定地成为两个进化支的共衍征，还需该属更多种类和居群的支持。在进化支II内部，花叶青木和青木聚为一支，亲缘关系较近。通过回溯花叶青木与青木的叶绿体全基因组序列，发现这两条叶绿体基因组序列，无论是基因数量还是基因位置，都保持高度一致，仅在基因间隔区存在4个碱基位点的变异。这也在一定程度上印证了 Ranney 的观点——青木中表达的斑点叶性状很可能是一个具有多重等位基因和数量遗传的核基因，而不是严格的母系遗传，可能有多个具有不同遗传模式的基因可以产生桃叶珊瑚属植物带斑叶的性状（Ranney，2018）。因此，“花叶”表现型，在桃叶珊瑚属类群中并不是一个稳定可靠的性状，基于“花叶”特征建立的变种或新种划分需格外谨慎。

向秋云在 FOC 中将峨眉桃叶珊瑚归并入原变种桃叶珊瑚(Xiang & Boufford, 20005)，本研究系统发育分析结果表明峨眉桃叶珊瑚与密花桃叶珊瑚关系更近，二者再与桃叶珊瑚构成姐妹群。Ranney 使用流式细胞仪对桃叶珊瑚属植物的相对基因组大小进行了研究，发现峨眉桃叶珊瑚的 1Cx 值（10.6 pg）显著高于包括桃叶珊瑚（7.37 pg）在内的其他种类（6.8~7.5 pg），这表明峨眉桃叶珊瑚经历了比其他类群更显著的倍性不变的基因组扩张（即基因组扩张非染色体加倍引起）(Ranney et al., 2018)；结合标本与野外宏观形态性状分析，三者构成了物种关系复杂的复合种群。因此，关于峨眉桃叶珊瑚的物种划分问题，还需要综合更多形态特征，尤其是细胞学证据，以及核基因数据的分析，特别是共线性单拷贝核基因，方能得到较清楚地解析。此外，一方面，桃叶珊瑚属植物很可能存在广泛的杂交和多倍化事件，另一方面，传统用以种间鉴别的叶形态特征在本属中变异极大，导致该属植物种间区别特征存在过渡变异和交叉类型，缺乏稳定可靠的区别特征，诸多种类难以准确辨别，再加上目前我们对于该属植物形态特征仍未有足够的把握，因此，对于物种鉴定不确切和物种划分明显不合理的种类，本研究未予囊括。对桃叶珊瑚属植物系统全面的属下分类系统的构建和属下种间关系的解决有赖于更广泛的采样，更多的性状证据（尤其是染色体数据），以及更丰富的具足够变异位点的单拷贝核基因和叶绿体基因组的联合分析来建树，来共同厘清桃叶珊瑚属的属下分类问题，为以后深入研究桃叶珊瑚属植物的药用和园艺用途等奠定分类学基础。

致谢 特别感谢河北师范大学石硕老师在黄杨叶丝缨花分子样品收集工作中给予的大力帮助。

参考文献：

ADEYEMO OA, AYODELE OO, AJISAFE MO, et al., 2021. Evaluation of dark jute SSR markers and morphological traits in genetic diversity assessment of jute mallow (*Corchorus olitorius* L.) cultivars[J]. S Afr J Bot, 137: 290-297.

AI TM, 2013. Medicinal Flora of China[M]. Beijing: Peking University Medicinal Press, 7: 279-287. [艾铁民，2013. 中国药用植物志[M]. 北京：北京大学医学出版社，7: 279-287.]

ALI AMIRYOUSEFI, JAAKKO HYVÖNEN, et al., 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes[J]. Bioinformatics, 17(34): 3030-3031.

Angiosperm Phylogeny Group, CHASE MW, CHRISTENHUSZ MJM, et al., 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV[J]. Bot J Linn Soc, 181(1): 1-20.

BASIC G, 2012. An integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. Bioinformatics, 28(12): 1647-1649.

BEIER S, THIEL T, MÜNCH T, et al., 2017. MISA-web: a web server for microsatellite prediction[J]. Bioinformatics, 33(16): 2583-2585.

BENSON G, 1999. Tandem repeats finder: a program to analyze DNA sequences[J]. Nucl Acid

Res, 27(2): 573-580.
CHEN C, CHEN H, ZHANG Y, et al., 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Mol Plant, 13(8): 1194-1202.
DARLING ACE, MAU B, BLATTNER FR, et al., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements[J]. Genome Res, 14(7): 1394-1403.
DOYLE JJ, DOYLE JL, 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue[J]. Phytochem Bull, 19: 11-15.
Editorial Committee of Chinese Materia Medica of the State Administration of Traditional Chinese Medicine, 1997. Chinese materia medica[M]. Shanghai: Science and Technology Press, 5: 734-737. [家中医药管理局中华本草编委会，1997. 中华本草[M]. 上海：科学技术出版社，5: 734-737.]
FRAZER KA, PACHTER L, POLIAKOV A, et al., 2004. VISTA: computational tools for comparative genomics[J]. Nucl Acid Res, 32(2): W273-W279.
HUANG Y, FAN L, HUANG J, et al., 2022. Plastome phylogenomics of *Aucuba* (Garryaceae)[J]. Front in Genet, 13: 753719.
JIA SN, ZHANG YM, ZHAO GF, et al., 2022. Comparative analysis of the chloroplast genomes of *Rhododendron capitalism*, *R. przewalskii*, and *Rhododendron* plants[J/OL]. Guihaia:1-17[2023-06-09].http://kns.cnki.net/kcms/detail/45.1134.Q.20220629.1301.002.html.. [贾守宁，张颖敏，赵国福，等，2022. 头花杜鹃、陇蜀杜鹃及杜鹃属植物叶绿体基因组比较分析[J/OL]. 广西植物 :1-17[2023-06-09].http://kns.cnki.net/kcms/detail/45.1134.Q.20220629.1301.002.html.]
JIANG JW, 2005. A Dictionary of Medicinal Plants[M]. Tianjin: Science and Technology: 92. [江纪武，2005. 药用植物辞典[M]. 天津：科学技术出版社：92.]
JIN JJ, YU WB, YANG JB, et al., 2020, GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes[J]. Genome Biol, 21(1): 1-31.
KAWABE A, MIYASHITA NT, 2003. Patterns of codon usage bias in three dicot and four monocot plant species[J]. Genes Genet Syst, 78(5): 343-352.
LOHSE M, DRECHSEL O, KAHLAU S, et al., 2013. OrganellarGenomeDRAW: a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets[J]. Nucl Acid Res, 41(W1): W575-W581.
MINH BQ, SCHMIDT HA, CHERNOMOR O, et al., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era[J]. Mol Biol Evol, 37(5): 1530-1534.
Nanjing University of traditional Chinese Medicine, 2006. Traditional Chinese medicine dictionary[M]. Shanghai: Science and Technology Press: 334. [南京中医药大学，2006. 中药大辞典[M]. 上海：科学技术出版社：334.]
PEDEN J, 2005. CodonW version 1.4. 2[CP]. Nottingham, UK: University of Nottingham.
POSADA D, 2008. jModelTest: phylogenetic model averaging[J]. Mol Biol Evol, 25(7): 1253-1256.
QU XJ, MOORE MJ, LI DZ, et al., 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes[J]. Plant Methods, 15(1): 1-12.
RANNEY TG, THOMASSON TH, NEILL K, et al., 2018. Ploidy, relative genome size, and inheritance of spotted foliage in *Aucuba* species (Garryaceae)[J]. HortScience, 53(9): 1271-1274.
ROMERO H, ZAVALA A, MUSTO H, 2000. Codon usage in *Chlamydia trachomatis* is the result

of strand-specific mutational biases and a complex pattern of selective forces[J]. Nucl Acid Res, 28(10): 2084-2090.

RONQUIST F, HUELSENBECK JP, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models[J]. Bioinformatics, 19(12): 1572-1574.

ROZAS J, FERRER-MATA A, SÁNCHEZ-DELBARRIO JC, et al., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets[J]. Mol Biol Evol, 34(12): 3299-3302.

ROZEWICKI J, LI S, AMADA KM, et al., 2019. MAFFT-DASH: integrated protein sequence and structural alignment[J]. Nucl Acid Res, 47(W1): W5-W10.

SHARP PM, Li WH, 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications[J]. Nucl Acid Res, 15(3): 1281-1295.

STEFAN K, CHOUDHURI JV, ENNO O, et al., 2001. REPuter: the manifold applications of repeat analysis on a genomic scale [J]. Nucl Acid Res, 29(22): 4633-4642.

WANG L, XING H, YUAN Y, et al., 2018. Genome-wide analysis of codon usage bias in four sequenced cotton species[J]. PLoS ONE, 13(3): e0194372.

XIANG QY, BOUFFORD DE, 2005. Flora of China[M]. Beijing: Science Press; St. Louis: Missouri Botanical Garden Press: 222-226.

XU C, CAI X, CHEN Q, et al., 2011. Factors affecting synonymous codon usage bias in chloroplast genome of oncidium gower ramsey[J]. Evol Proteins, 7: EBO. S8092.

ZENG X, GUO F, OUYANG D, 2020. A review of the pharmacology and toxicology of aucubin[J]. Fitoterapia, 140: 104443.